TITLE:

相手の位置を不完備情報とする移動目標捜索ゲーム (不確実性下における意思決定問題)

AUTHOR(S):

樋口, 英樹; 宝崎, 隆祐

CITATION:

樋口, 英樹 ...[et al]. 相手の位置を不完備情報とする移動目標捜索ゲーム (不確実性下における意思決定問題). 数理解析研究所講究録 2011, 1734: 33-40

ISSUE DATE:

2011-03

URL:

http://hdl.handle.net/2433/170781

RIGHT:

京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

数理解析研究所講究録
第 1734 巻 2011 年 33-40

相手の位置を不完備情報とする移動目標捜索ゲーム

防衛大学校・理工学研究科　樋口　英樹(Hideki Higuchi)
Graduate School of Science and Engineering,
National Defense Academy
防衛大学校・情報工学科　宝崎　隆祐(Ryusuke Hohzaki)
Department of Computer Science,
National Defense Academy

1　はじめに

第2次大戦時の潜水艦に対する最適な捜索方法の研究を起源とする捜索問題は, 1980年代以降, 捜索者と逃避者をプレイヤーとする移動目標捜索ゲームにより主として定式化されている. このゲームでは, 各プレイヤーは互いに相手の動きを予想して自らの動きを決定する. したがって各プレイヤーが持つ相手の存在範囲に関する知識が, 最適な捜索・逃避運動を大きく変化させる.

この知識に関して従来研究では, 両プレイヤーとも常に相手の位置を全く知らない([1,2]), または完全に把握している([3-5])と仮定されている. つまり相手の位置に関する知識はゲームのプレイ中変化せず, プレイヤーは位置把握度を考慮せずに意思決定を行うことが可能であった.

しかし現実においては, 相手の位置の把握度合いは通常両プレイヤーの行動により変化するため, 両プレイヤーはその影響を考慮して行動の決定を行うと考えられる. またこの考慮が, 捜索者と逃避者の行動決定にジレンマを引き起こすと言える. 例えば暗闇の中での捜索において, 捜索者は単純に探知のみを考えた場合は懐中電灯を使用すべきであるが, しかし自らの位置を逃避者にさとられないためには, 目や耳等の能力の低いセンサーによる捜索が適切であろう.

このような位置把握度の影響やジレンマを考慮した研究として[6]や[7]が存在する. しかし前者は, 逃避者の先制探知による捜索者の探知能力の変化をパラメータで表現しており, その変化の程度が事前に把握不可能な一般の場合に適用できない. 一方後者も, 捜索空間が円環と非常に限定されており, また逃避者の行動も考慮されていない.

そこで本研究では, 一般的な捜索状況における位置把握度の考慮による最適な捜索・逃避運動の変化, 及びジレンマを持った行動手段の最適な選択方法を明らかにするため, 相手の位置を不完備情報とする移動目標捜索ゲームをモデル化した. そして, 位置把握度に応じたプレイヤーの意思決定を有限状態制御子(Finite State Controller, FSC)で表現し, 望ましい FSC を, Sampled Saddle Point(SSP)法を用いて確率的に評価した.

本論文の構成は以下のとおりである. まず第 2 節で, 相手の位置を不完備情報とする移動目標捜索ゲームの基本的な仮定を説明し, 定式化を行う. そして第 3 節で FSC による戦略の表現法, 第 4 節で FSC を戦略とする大規模利得行列の確率的な解の保証の得方を説明する. 第 5 節では, 簡単な数値例により捜索者及び逃避者の戦略の評価を行い, 最後に第 6 節で今後の課題を述べる.

2　相手の位置を不完備情報とする移動目標捜索ゲーム

ここでは, 以下のような前提を持った捜索者 S と逃避者 E との間の捜索ゲーム Γ を考える.

(A1)　捜索空間は, 2 次元平面を n 個のセルに分割した離散地理空間 $\mathbf{N}=\{1,2,\cdots,n\}$ とする.

(A2)　プレイヤーの集合を $\mathbf{K}=\{S,E\}$ とし, 捜索者 S は逃避者 E を捜索し, E は S から逃避する.

(A3) プレイヤー$k \in \mathbf{K}$は，現在位置$\mathbf{x}'_k \in \mathbf{N}$からの移動可能範囲$\mathbf{M}_k(\mathbf{x}'_k) \subseteq \mathbf{N}$と，行動手段の集合$\mathbf{B}_k = \{1,2,\cdots,m_k\}$を持ち，位置$\mathbf{x}_k \in \mathbf{M}_k(\mathbf{x}'_k)$と行動手段$b_k \in \mathbf{B}_k$の組からなる行動$\mathbf{a}_k = (\mathbf{x}_k, b_k) \in \mathbf{A}_k = \mathbf{N} \times \mathbf{B}_k$を各離散時刻で選択して捜索又は逃避を行う．ここで各セルからの移動可能先は縦横の隣接セルのみとし，$\mathbf{M}_k(\mathbf{x}'_k)$は単位時間以内にプレイヤーの速度で移動可能なセルの範囲を表すものとする．

(A4) 捜索者と逃避者が行動の組$\mathbf{a} = (\mathbf{a}_S, \mathbf{a}_E) \in \mathbf{A} = \mathbf{A}_S \times \mathbf{A}_E$を選択した時，各プレイヤー は互いに独立に確率$\eta_S(\mathbf{a}), \eta_E(\mathbf{a})$で相手を探知し，行動手段$b_k$に依存する精度$\sigma_k(b_k)$で相手の存在範囲を認識する．ここで精度は，認識する存在範囲の誤差の大きさを表すものとする．

(A5) プレイヤーkは相手の行動と探知の有無を直接把握できず，自らの行動と探知の有無に基づき相手の位置を信念$\rho_k \subseteq \Delta(\mathbf{N})$で予想する．ただし，$\Delta(\mathbf{N})$は$\mathbf{N}$上の確率分布を表す．

(A6) プレイヤーの初期位置の組は，初期信念の組$\rho^0 = (\rho_S^0, \rho_E^0)$と整合的に決定される．

(A7) 両プレイヤーは離散時刻で行動を選択し，捜索者が探知した場合，捜索者は利得 1 を，逃避者は利得-1 を得る．探知しない場合，両プレイヤーは利得 0 を得る．

(A8) ゲームは捜索者の探知のみにより終了する．

(A9) プレイヤーは自らの信念の下での割引総利得最大化を目的に行動を選択する．

このゲームは，(A2)及び(A5)から，プレイヤーの位置の組を状態とする部分観測可能な確率過程ゲーム(Partially Observable Stochastic Game, POSG)として，以下のように定義できる．

定義 2.1 位置の組を状態とする POSG 形式の捜索ゲームΓは次のとおり定義される．

$$\Gamma = \left\langle \mathbf{K}, \mathbf{S}, \rho^0, \{\mathbf{A}_k\}_{k \in K}, \{\mathbf{O}_k\}_{k \in K}, P_{STEP}, P_{SIG}, \{R_k\}_{k \in K} \right\rangle \tag{1}$$

ここで，

(1) $\mathbf{K}$はプレイヤーの集合

(2) $\mathbf{S} = \mathbf{N} \times \mathbf{N}$は状態の有限集合であり，プレイヤーの位置の組

(3) $\rho^0 \in \Delta(\mathbf{S})$は初期の状態分布

(4) $\mathbf{A}_k$，$k \in \mathbf{K}$はプレイヤーkの行動集合

(5) $\mathbf{O}_S = \{0\}, \mathbf{O}_E = \{0,1\}, \mathbf{O} = \mathbf{O}_S \times \mathbf{O}_E$は捜索者$S$と逃避者$E$が行動選択後に受信するシグナルの集合，並びにシグナルの集合の組であり，0は非探知，1は探知を表す．

(6) $P_{STEP} : \mathbf{S} \times \mathbf{A} \to \Delta(\mathbf{S})$は行動選択後の状態遷移確率を表し，下式で与えられる．

$$P_{STEP}(\mathbf{y}'|\mathbf{y},\mathbf{a}) = \delta_{\mathbf{y}'\mathbf{x}}, \quad \mathbf{y},\mathbf{y}' \in \mathbf{S}, \quad \mathbf{a} = ((\mathbf{x}_S, b_S),(\mathbf{x}_E, b_E)) \in \mathbf{A}, \quad \mathbf{x} = (\mathbf{x}_S, \mathbf{x}_E) \in \mathbf{S} \tag{2}$$

ただし，$\delta_{\mathbf{y}'\mathbf{x}}$はディラックの$\delta$関数を表す．

(7) $P_{SIG} : \mathbf{A} \to \Delta(\mathbf{O})$は行動選択後のシグナルの発生確率を表し，下式で与えられる．

$$P_{SIG}(\mathbf{o}|\mathbf{a}) = \begin{cases} (1-\eta_S(\mathbf{a}))(1-\eta_E(\mathbf{a})) & if\ \mathbf{o} = (0,0) \\ (1-\eta_S(\mathbf{a}))\eta_E(\mathbf{a}) & if\ \mathbf{o} = (0,1) \end{cases}, \quad \mathbf{o} \in \mathbf{O}, \quad \mathbf{a} \in \mathbf{A} \tag{3}$$

(8) $R_k : \mathbf{A} \to \Re$はプレイヤーkの利得関数を表し，下式で与えられる．

$$R_S(\mathbf{a}) = \eta_S(\mathbf{a}), \quad \mathbf{a} \in \mathbf{A} \tag{4}$$

$$R_E(\mathbf{a}) = -\eta_S(\mathbf{a}), \quad \mathbf{a} \in \mathbf{A} \tag{5}$$

3 位置把握度を考慮した戦略の表現

2 節で定義したΓに対するプレイヤーの戦略の表現法として，本研究では FSC[8]を使用する．FSC は，POSG の一方的意思決定問題版である部分観測可能なマルコフ意思決定過程(Partially Observable Markov Decision Process, POMDP)で使用される意思決定の表現方法

である. FSCではノードと呼ばれる内部状態を考え，そのノード毎に選択する行動を指定する．そしてノード間は受信するシグナルに応じて推移する．このようにして，シグナルに応じたプレイヤーの無限期間の行動の使い分けを表現する．

このFSCは，通常のPOMDPでは[8]のように動的計画法を用いてその最適な構造を獲得する．しかし作成されるFSCの構造は一般に複雑になり，そのFSCが表現する意思決定の内容の解釈は困難である．またそもそもPOSGに対して，このFSCの構造獲得アルゴリズムは開発されていない．そこで本研究では，FSCの構造をあらかじめ固定して考える．

FSCの構造を固定する場合，その妥当な構造とその下での複雑な行動の実現法を検討する必要がある．まず前者については，プレイヤーの位置把握度に応じた意思決定を反映した単純な構造を考える．そこでまず捜索者については，常に相手を探知しない状況で捜索を継続することから，ノードが1つのFSCとする．一方逃避者は，探知の有無によって捜索者の存在範囲に関する認識が変化し，この認識に基づき行動を変化させると考えられる．そこで認識する存在範囲の誤差の大きさをノードとするFSCを考え，ノード間は探知・非探知に応じた誤差の大きさの変化により推移するものとする．また誤差が一定以上となった時点で逃避者は非探知状態になったと認識するものとする．このようなFSCの具体例を図1に示す．

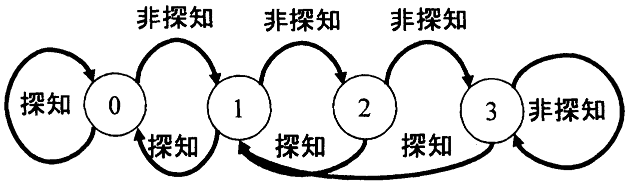


図1　逃避者のFSCの例（捜索者の速度を1，逃避者の手段の精度を0と1，存在範囲の大きさの最大値を3とした場合）． ノードの数値は逃避者が認識する捜索者の存在範囲の誤差の大きさを示す．“探知”，“非探知”は逃避者が受信するシグナルを表し，“探知”の場合はその探知手段の精度，“非探知”の場合は捜索者の速度により現在ノードが推移する．誤差3で逃避者は非探知になったと認識するとし，対応するノードを非探知ノードと呼ぶ．

一方後者のために，本研究では移動と手段からなる行動を生成するルールを各ノードに指定する．使用するルールは，移動は表1のような一般的な捜索・逃避パターンを，手段は表2のような特定の手段の選択法を考える．なお表1で，“位置情報”は移動時の位置情報の利用の有無を，“網羅性”は捜索空間全体の網羅の有無を表す．

表1　プレイヤーの移動ルール．

No	位置情報	網羅性	ルール名	移動方法	対象
1	非利用	非網羅	RND	移動可能範囲内のセルをランダムに選択	両方
2			停止	現在地に停止	両方
3		網羅	一様拡散	捜索空間全体に一様に拡散	両方
4			手段拡散	特定の手段の使用回数を空間全体で拡散	捜索者
5			平行捜索	捜索空間に沿って平行に移動	捜索者
6	利用	-	近接・離隔	目標予想位置に近接・離隔するセルへ移動	逃避者

表2　プレイヤーの手段選択ルール．

No	ルール名	手段選択方法	戦略要素	対象
1	一定間隔	指定された時間間隔	時間間隔	捜索者
2	指数間隔	指数分布に従う時間間隔	平均時間間隔	
3	一定確率	毎時指定された確率	確率	
4	指定	指定された手段を常に選択	-	逃避者

4 大規模な行列ゲームに対する確率的な解の保証

前節で説明したFSCを戦略とする利得行列を作成し，そのミニマックス解を求めれば各プレイヤーの最適FSCを求めることができる．しかし，FSCの組の利得の計算にはモンテカルロシミュレーションが必要であり，また逃避者のFSCの数はルール数に対するノード数のべき乗と膨大になることから，厳密なミニマックス解を求めることは事実上不可能である．そこで本研究では，SSP法[9]を用いて，大規模な2人ゼロ和行列ゲームの解に対する確率的な保証値を求め，その保証値の大小により戦略の優劣を検討する．

SSP法は，サンプリングした利得行列により各プレイヤーの解を求め，その解に対する保証値を確率的に評価する．この解をサンプルセキュリティポリシー(ssp)，保証値をサンプル保証値と呼ぶ．このとき，$M\times N$ の大規模な利得行列 F に対し，サンプル保証値 $\overline{V}(F_1)$ を持った最小化プレイヤー P_1 の ssp y^*,サンプル保証値 $\underline{V}(F_2)$ を持った最大化プレイヤー P_2 の ssp z^* を求めるアルゴリズムは以下のとおりである。なお， $\Lambda^{k\times l}$ は $k\times l$ 左 01 確率行列を表す．

Algorithm1: Sampled Saddle Point Algorithm

1. P_1 に対し，ランダムサンプリングにより確率行列 $\Gamma_1\in\Lambda^{M\times m_1}$ 及び $\Pi_1\in\Lambda^{N\times n_1}$ を作成する．
2. F の $m_1\times n_1$ 部分行列 $F_1=\Gamma_1' F\Pi_1$ を作成する．
3. サンプリングした m_1 個の純戦略上の P_1 の混合戦略の集合を S_{m_1} , n_1 個の純戦略上の P_2 の混合戦略の集合を S_{n_1} とし，ssp y^* を $y_1^*\equiv\arg\min_{y_1\in S_{m_1}}\max_{z\in S_{n_1}} y_1' F_1 z$ として $y^*=\Gamma_1 y_1^*$ により求める．
4. y^* に対するサンプル保証値 $\overline{V}(F_1)$ を $\overline{V}(F_1)=\max_{z\in S_{n_1}} y_1^{*\prime} F_1 z$ により求める．
5. P_2 に対しても，1〜4 で添え字を 1→2，max ⇔ min とすることで z^* 及び $\underline{V}(F_2)$ が求まる．

このsspに対する保証は以下のように定義され，またその保証を与えるサンプルサイズは以下の定理で与えられている（以降全て P_1 についての結果のみ示す）．

定義 4.1 サンプル保証値 $\overline{V}(F_1)$ を持つ ssp y^* に対し，

$$P_{\Gamma_2,\Pi_2}\left(y^{*\prime} F z^*\le\overline{V}(F_1)+\varepsilon\,\middle|\,y^*,\overline{V}(F_1)\right)\ge 1-\delta \tag{6}$$

が成立するとき，「y^* は信頼度 $1-\delta$ で ε 安全である」という．ここで P_{Γ_2,Π_2} は Γ_2 及び Π_2 に関するサンプリングの確率を表す．

定理 4.1 ([9]) $\Gamma_1,\Pi_1,\Gamma_2,\Pi_2$ が統計的に独立で，かつ Π_1 と Π_2 が同一の分布に従うとする．このとき，ある $\gamma\in(0,1)$ と $\overline{n}_2\ge n_2$ に対し，n_1 が

$$n_1=\left\lceil\frac{1}{\delta}\left(\ln\frac{1}{\gamma}+m_1+\sqrt{2m_1\ln\frac{1}{\gamma}}\right)\right\rceil\overline{n}_2 \tag{7}$$

を満足するとき，$1-\gamma$ 以上の確率で，$\overline{V}(F_1)$ を持つ ssp y^* は信頼度 $1-\delta$ で $\varepsilon=0$ 安全である．

(7)はサンプル保証値を得る前に必要なサンプルサイズの見積りを表すが，事前に(7)を満足するサンプルサイズを確保できずに ssp y^* とサンプル保証値 $\overline{V}(F_1)$ を求めた後，追加のサンプリングで y^* の保証を事後的に得ることが可能である．これを行うには，Algorithm1 で y^* を求めた後，Π_1 と独立に同一の分布で $\overline{\Pi}_1\in\Lambda^{N\times k_1}$ を作成し，以下の値を計算する．

$$\overline{v}=\max_{j\in\{1,\cdots,k_1\}} y^{*\prime} F\overline{\Pi}_1 e_j \tag{8}$$

ここでe_jは単位ベクトルを表す．この$\bar{v}$による事後の保証を得るための追加のサンプルサイズk_1が，以下の定理で与えられている．

定理 4.2 ([9]) $\Gamma_1, \Pi_1, \Gamma_2, \Pi_2$が統計的に独立で，かつ$\Pi_1$と$\Pi_2$が同一の分布に従うとする．このとき，ある$\gamma \in (0,1)$と$\bar{n}_2 \geq n_2$に対し，$k_1$が

$$k_1 = \left\lceil \frac{1}{\delta} \ln \frac{1}{\gamma} \right\rceil \bar{n}_2 \tag{9}$$

を満足するとき，$\varepsilon \geq \bar{v} - \bar{V}(F_1)$を満足する任意の$\varepsilon$に対し，$1-\gamma$以上の確率で，$\bar{V}(F_1)$を持つ$y^*$は信頼度$1-\delta$で$\varepsilon$安全である．

以上より，所定のサイズのサンプリングで，信頼度を持った解と保証値を得ることができる．

5 数値例

本節では，第 4 節の SSP 法を用いて，簡単な数値例により捜索者及び逃避者の望ましい FSC の評価を行う．以降の全ての計算で使用する基本的な設定を表 3 に示す．

表 3　数値計算での基本的な設定

項目	設定値	項目	設定値
捜索空間N	9×9	初期信念ρ_0	両プレイヤーとも一様
割引因子	0.99	探知関数η	両プレイヤーともユークリッド距離に対する逆 3 乗関数
モンテカルロ試行回数	1,000		
存在圏の大きさの最大値	4	速度	両プレイヤーとも 1

また捜索者と逃避者の行動手段については，各々ジレンマがある 2 つの手段を考える．捜索者は探知重視の手段 A と被探知回避重視の手段 P，逃避者は逃避重視の手段 L と位置把握重視の手段 W の 2 つを考え，それらの組について，探知距離及び探知精度を表 4 のように設定する．なお，探知距離は逆 3 乗関数から求まる有効探知距離を，探知精度は探知時に得られる存在圏の標準偏差の大きさを表す．表 4 の設定を，以下では基準ケースと呼ぶ．

表 4　捜索者と逃避者の手段の組毎のパラメータ設定

(ア) 捜索者の探知距離

	L	W
A	1	1
P	0.5	0.5

(イ) 逃避者の探知距離

	L	W
A	2	0.5
P	1	0.5

(ウ) 逃避者の探知精度

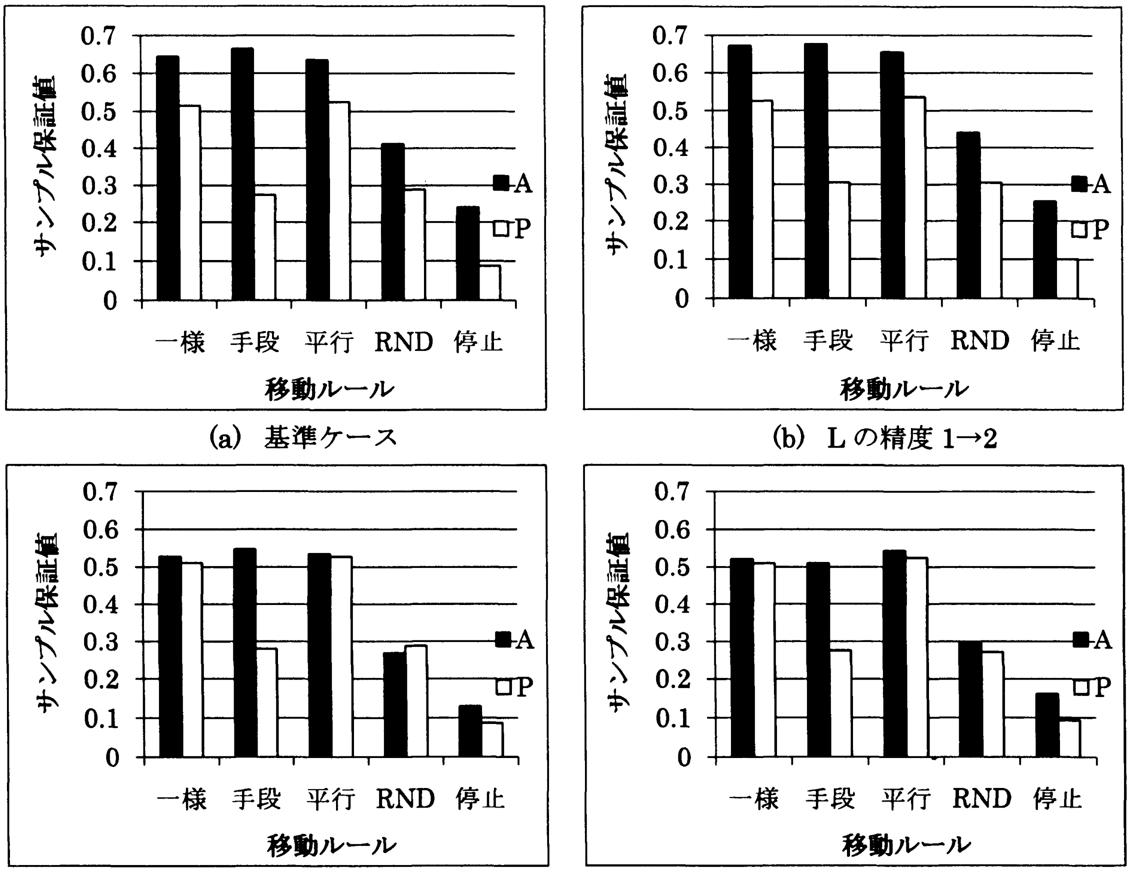
手段	精度
L	1
W	0

5.1. 捜索者の FSC の評価

まず捜索者の FSC の評価を行う．評価対象の FSC は，比較を容易にするため，手段選択を A か P の 2 通りのみとし，移動ルールは表 1 の No.1 から 5 を考える．そして SSP の計算では，評価対象の FSC を 1 つずつ考え，つまり$m_1 = m_2 = 1$とし，それに対し$n_1 = 5$, $n_2 = 10$, $\delta = \gamma = 0.05$として，一様サンプリングにより事後的な保証値を計算した．結果を図 2 に示す．

図 2 でまず移動ルールについてみていくと，(a)～(d)の全てのケースで，RND や停止といった捜索空間を網羅しない移動方法が，一様拡散や平行捜索のような捜索空間を網羅する移動方法に劣っていることが分かる．一方網羅型の捜索の間では，従来最適解である一様拡散も，行動が読まれ易い平行捜索もサンプル保証値に大きな差はないことが分かる．このことから，捜索者は逃避者に動きを読まれるかどうかを意識することなく，捜索空間全体を網羅するように運動するのが望ましいということがわかる．

一方，手段の選択についてみていくと，(a)及びLの精度を下げた(b)では，探知重視の行動Aが望ましいのが分かる．しかし(c)と(d)のように，Aの探知能力または被探知回避性を下げると，Aと被探知回避重視の行動Pの間にほぼ差がなくなることが分かる．このことから捜索者は，探知あるいは被探知に関する手段間の能力の差によって，探知重視と被探知回避重視の手段間のトレードオフを判断できることが分かる．



(a) 基準ケース (b) Lの精度1→2

(c) Aの探知距離1→0.75 (d) Aに対するLの探知距離2→4

図2 捜索者の95%サンプル保証値の比較． 各グラフは基準ケースに対し，(b) Lの精度を1→2，(c) Aの探知距離を1→0.75，(d) Aに対するLの探知距離を2→4にした場合を示している．

5.2. 逃避者のFSCの評価

次に逃避者のFSCの評価を行う．評価対象とするFSCは，移動については，逃避者が持つ捜索者の位置情報の利用方法で表5のように3通り考え，また手段については，精度の追求の観点で表6のように3通り考える．

表5 評価対象FSCの移動ルールの設定

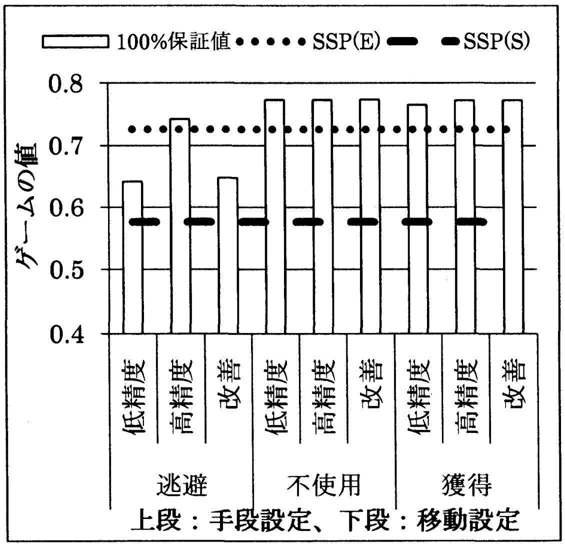
設定名	FSCにおける移動ルールの設定
逃避	非探知ノード以外は全て「離隔」とし，非探知ノードでは「停止」．
不使用	全てのノードで「一様拡散」．
獲得	Lの精度以下のノードでは「離隔」，非探知ノードでは「一様拡散」，その他のノードでは「近接」．

表6 評価対象FSCの手段設定

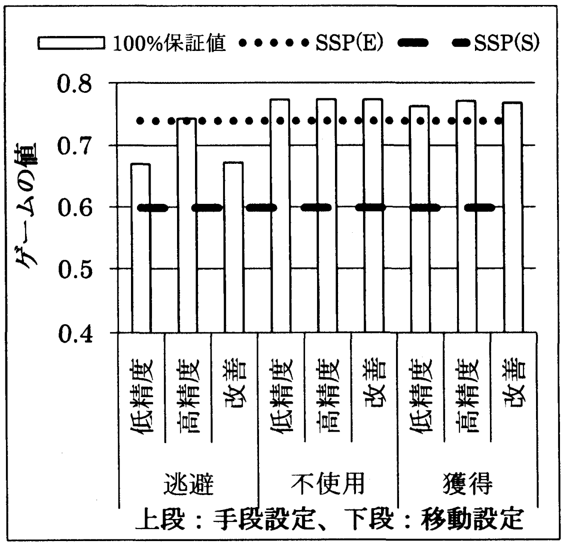
設定名	FSCにおける手段の設定
低精度	全てのノードでL.
高精度	全てのノードでW.
改善	Lの精度以下のノードではW，その他のノードではL.

評価にあたっては，まず捜索者の全FSCを考え，それらに対する評価対象のFSCのミニマ

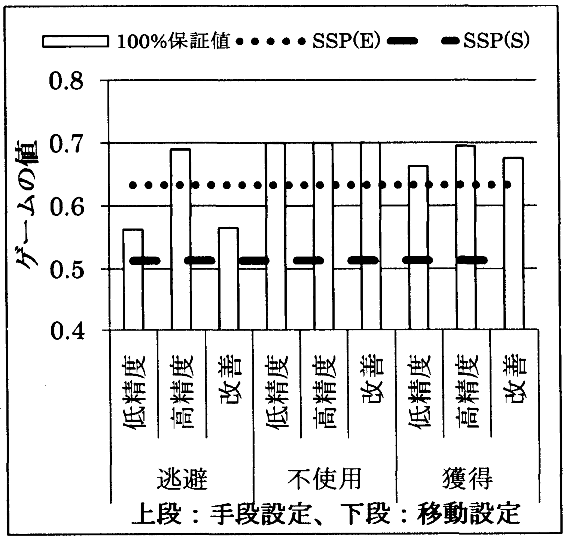
ックス値を求める．この値は，捜索者が逃避者の FSC を知っている場合に逃避者が確保可能な最低の値であり，評価対象の FSC の 100%保証値である．次に，$m_1 = 3, n_1 = 10, n_2 = 3, m_2 = 10$ $\delta = \gamma = 0.05$ として一様サンプリングにより事後的な SSP を行い，捜索者及び逃避者のサンプル保証値を得る．これを 100%保証値と比較することで，評価対象の FSC のよさの程度が評価可能である．以上の計算結果を図 3 に示す．



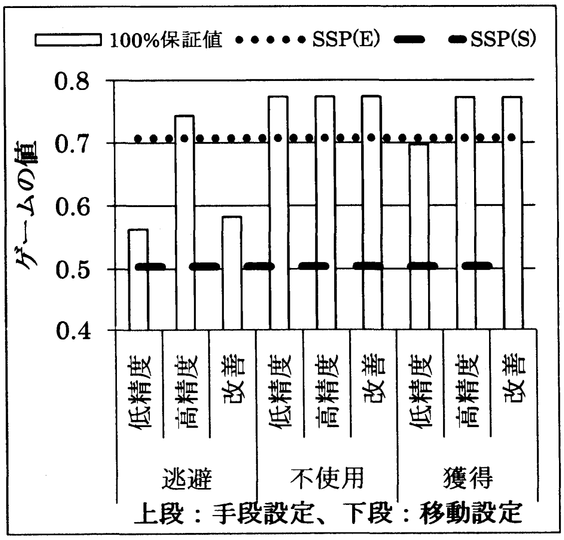
(a) 基準ケース



(b) E の精度 1→2



(c) A の探知距離 1→0.75



(d) A に対する E の探知距離 2→4

図 3　逃避者の 100%保証値と 95%サンプル保証値． 各グラフの設定は図 2 と同じである．また SSP(E)は逃避者の，SSP(S)は捜索者の 95%サンプル保証値を表す．

図 3 で 4 つのグラフは全て同じ傾向を示している．まず移動ルールは，位置情報を逃避に使用する FSC の 100%保証値だけが，捜索者及び逃避者のサンプル保証値の間にあり，当該 FSC がランダムにサンプリングした FSC よりも明らかに良いことが分かる．一方位置情報を逃避に使用しない，あるいは情報獲得に使用する FSC は，サンプリングで作成した FSC の保証値と同じかそれより悪い値しか保証されておらず，逃避者にとって望ましくない逃避法であると言える．

次に手段の選択については，高精度の手段のみを選択する FSC は逃避者にとって明らかに望ま

しくないのが分かる．一方低精度の手段のみ，あるいは精度を改善する手段を選択する FSC の差はほとんどなく，逃避者は精度を考慮せず，探知能力のみを考慮して手段を選択すればよいのが分かる．図 2 で，捜索者は探知重視の優位性が薄れるにつれて被探知回避の手段の選択が望ましくなったが，その変化に対応した逃避者側の手段選択の変化は図 3 では見られなかった．

6　おわりに

本研究では，相手プレイヤーの位置の把握度を考慮した捜索ゲームを定式化し，位置把握度に応じた捜索・逃避方法及び捜索・逃避手段の選択について論じた．その結果，捜索者は探知性と被探知性に応じて捜索手段を使い分けつつ，捜索空間を網羅する捜索が，逃避者は精度を考慮せずに捜索者の探知に努め，得られた情報を使って逃避を行うのが望ましいと分かった．

今後は 2 種類以上の手段に関する分析や，行動に関する制約を有する手段の追加等のモデルの高度化を図るとともに，これらの高度化に対応可能なより効率的なサンプリングによる解の評価方法を検討する予定である．

参考文献

[1] J. Eagle and A. Washburn, "Cumulative search-evasion games," *Naval Research Logistics*, vol. 38, 1991, pp. 495-510.

[2] R. Hohzaki, "Search allocation game," *European Journal of Operational Research*, vol. 172, Jul. 2006, pp. 101-119.

[3] R. Hohzaki, "A multi-stage search allocation game," *Journal of the Operations Research Society of Japan*, vol. 50, 2007, pp. 178-200.

[4] L. Thomas and A. Washburn, "Dynamic search games," *Operations Research*, vol. 39, 1991, pp. 415-422.

[5] A. Washburn, "Search-evasion game in a fixed region," *Operations Research*, vol. 28, 1980, pp. 1290-1298.

[6] Koji Iida, "Hide-and-search game taking account of forestalling detection by the target," *Mathematica Japonica*, vol. 44, 1996, pp. 245-260.

[7] G. Olsder, "About when to use the searchlight," *Journal of Mathematical Analysis and Applications*, vol. 136, Dec. 1988, pp. 466-478.

[8] E. Hansen, "Solving POMDPs by searching in policy space," *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 1998, pp. 211-219.

[9] S.D. Bopardikar, A. Borri, J.P. Hespanha, M. Prandini, and M.D. Di Benedetto, "Randomized sampling for large zero-sum games," *Technical Report*, CA, USA 93106: Department of Electrical and Computer Engineering, University of California Santa Barbara, 2010, pp. 1-20.